μ measure on $\mathcal{X}$ w.r.t. metric D. $\mathcal{X} \subseteq \mathbb{R}^d$

$X_1, \ldots, X_n \overset{iid}{\sim} \mu$, empirical measure:

$$\mu_n = \frac{1}{n} \sum_{i=1}^{n} \delta_{X_i}.$$

$W_p(\mu_n, \mu)$?

$$W_\infty(\mu, \nu) = \sup_{p \geq 1} W_p(\mu, \nu)$$

$$\underline{W_\infty(\mu_n, \mu) \not\longrightarrow 0}$$

$$W_p(\mu_n, \mu) \longrightarrow 0 \quad a.s.$$

<u>Goal</u>: Bound

$$\mathbb{P}\big(\underbrace{W_p(\mu_n, \mu)}_{\text{"Empirical Wasserstein Distance"}} \geq u \big) \leq \ldots, \quad u > 0.$$

<u>Outline</u>:

(1) Bound

$$\mathbb{P}\big(|W_p(\mu_n, \mu) - \mathbb{E} W_p(\mu_n, \mu)| \geq u \big)$$

② Bound $\mathbb{E}\, W_p(\mu_n, \mu)$.

<u>Part ①</u> : Deviations from the Mean.

<u>Fact 1</u> : $\mu \in T_p(\sigma^2)$

$$\Longrightarrow \mu^{\otimes n} \in T_p(\underbrace{n^{\frac{2}{p}-1}\sigma^2}_{}) \quad \forall p \geq 1$$

$$D_p(x,y) = \left(\sum_{i=1}^{n} D^p(x_i, y_i) \right)^{1/p}$$

$x, y \in \mathcal{X}^n$.

<u>Fact 2</u> : If $\nu \in \mathcal{P}(\mathcal{Y})$ and $\nu \in T_1(\sigma^2)$, then $f(Y)$, where $Y \sim \nu$, is σ^2-sub-Gaussian for all 1-Lipschitz maps $f : \mathcal{Y} \to \mathbb{R}$.

<u>Theorem</u> : Suppose $\mu \in T_p(\sigma^2)$. Then $W_p(\mu_n, \mu)$ is σ^2/n-sub-Gaussian, $\forall n \geq 1$.

In particular,

$$(*) \quad \mathbb{P}\left(|W_p(\mu_n,\mu) - \mathbb{E} W_p(\mu_n,\mu)| \geq u \right) \leq 2 e^{-\frac{nu^2}{2\sigma^2}}$$

$\forall u > 0.$

$$\left(\text{i.e } |W_p(\mu_n,\mu) - \mathbb{E} W_p(\mu_n,\mu)| \lesssim \sqrt{\frac{\sigma^2}{n} \log(1/\delta)} \ \text{w.p.al. } 1-\delta \right)$$

cf. Niles - Weed & Rigollet (2019).

Proof.

Define :

$$f : (x_1, \ldots, x_n) \in \mathcal{X}^n \longmapsto W_p\left(\frac{1}{n} \sum \delta_{x_i}, \mu\right)$$

$$|f(x_1, \ldots, x_n) - f(y_1, \ldots, y_n)|$$

$$\leq W_p\left(\frac{1}{n} \sum_{i=1}^n \delta_{x_i}, \frac{1}{n} \sum_{i=1}^n \delta_{y_i} \right)$$

$$= \min_{\tau \in S_n} \left[\frac{1}{n} \sum_{i=1}^n D(x_i, y_{\tau(i)})^p \right]^{1/p}$$

$$\leq \left(\frac{1}{n} \sum_{i=1}^n D^p(x_i, y_i) \right)^{1/p}.$$

$$= n^{-\frac{1}{p}} D_p(x, y).$$

$\Longrightarrow f$ is $n^{-\frac{1}{p}}$ Lip.

$\Longrightarrow n^{1/p} f$ is 1-Lip.

But $\mu^{\otimes n} \in T_p\left(n^{\frac{2}{p}-1}\sigma^2\right) \Longrightarrow \mu^{\otimes n} \in T_1\left(\underbrace{n^{\frac{2}{p}-1}\sigma^2}\right)$

$\Longrightarrow n^{\frac{1}{p}} f(X_1, \dots, X_n)$ is $n^{\frac{2}{p}-1}\sigma^2$-sub-Gaussian

$\Longrightarrow \underline{f(X_1, \dots, X_n)}$ is $\dfrac{\sigma^2}{n}$-sub-Gaussian $\square$

② Bounding $\mathbb{E} W_p^p(\mu_n, \mu)$. ($\mathcal{X}$ compact)

Let $\mathcal{F}$ be a class of functions on $\mathcal{X}$.

Define: $d_{\mathcal{F}}(\mu, \nu) = \sup_{f \in \mathcal{F}} \int f \, d(\mu - \nu)$.

e.g.: $d_{\mathcal{F}}(\mu, \nu) = W_1(\mu, \nu)$ if

$$\mathcal{F} = \mathcal{F}_{Lip} = \{ f : \|f\|_{Lip} \leq 1 \},$$

$$d_{\mathcal{F}}(\mu_n, \mu) = \sup_{f \in \mathcal{F}} \frac{1}{n} \left\{ \sum_{i=1}^{n} \left[f(X_i) - \mathbb{E} f(X_i) \right] \right\}$$

__Definition__: Given a metric space $(\mathcal{Y}, \rho)$ and $\varepsilon > 0$, an ε-cover is a collection $\theta_1, \dots, \theta_N \in \mathcal{Y}$ such that

$$\forall \theta \in \mathcal{Y}, \quad \exists\, i \in \{1, \dots, N\} : \rho(\theta_i, \theta) \leq \varepsilon.$$

If such N exists, the smallest such is called the ε-covering number of $\mathcal{Y}$, and we write: $N = \underline{N(\varepsilon, \mathcal{Y}, \rho)}$.

__e.g__: $N(\varepsilon, [0,1]^d, \|\cdot\|) \asymp \left(1/\varepsilon\right)^d$

__Theorem__ (Dudley's Chaining)

If $\mathcal{F}$ is uniformly bounded by $b > 0$, then for all $\tau > 0$:

$$\rightarrow \mathbb{E}\left[d_{\mathcal{F}}(\mu_n, \mu)\right] \lesssim \tau + \frac{1}{\sqrt{n}} \int_{\tau}^{2b} \sqrt{\log N(\varepsilon, \mathcal{F}, L^\infty)}\, d\varepsilon$$

__e.g__: $N(\varepsilon, \mathcal{F}_{Lip}, L^\infty) \lesssim \left(\frac{1}{\varepsilon}\right)^{N(\varepsilon, \mathcal{X}, D)}$

If $\mathcal{X} = [0,1]^d$, then

$$\log N(\varepsilon, \mathcal{F}_{Lip}, L^\infty) \leq \log(1/\varepsilon)\left(\frac{1}{\varepsilon}\right)^d$$

$$\Longrightarrow \mathbb{E}[W_1(\mu_n,\mu)] \lesssim n^{-1/d} f(n)$$

<u>Proof Idea</u> : Let $\tau > 0$. Let $f_1, \ldots, f_N$ be a τ-cover, where $N = N(\tau, \mathcal{F}, L^\infty)$.

$\forall f \in \mathcal{F}$, $\exists i : L^\infty(f, f_i) \leq \tau$,

$$\int f \, d(\mu_n - \mu)$$

$$= \underbrace{\int (f - f_i)\, d(\mu_n - \mu)}_{\lesssim \tau} + \int f_i \, d(\mu_n - \mu)$$

$$\underline{\sup_{f \in \mathcal{F}}} \int f \, d(\mu_n - \mu) \lesssim \tau + \underbrace{\max_{1 \leq i \leq N} \int f_i \, \underline{d(\mu_n - \mu)}}$$

$$\rightarrow \mathbb{E}[\cdot] \lesssim \sqrt{\frac{\log N}{n}}$$

Better Approach.

In what follows, write

$$N_\varepsilon(S) = N(\varepsilon, S, D)$$

• Dyadic Partition: A collection $\{Q_k\}_{k=1}^{k^*}$ of partitions of $\mathcal{X}$ s.t:

(i) $\mathcal{X} = \bigcup_{S \in Q_k} S \quad \forall k = 1, \dots, k^*$

(ii) $\mathrm{diam}(S) \leq 3^{-k} \quad \forall S \in Q_k$

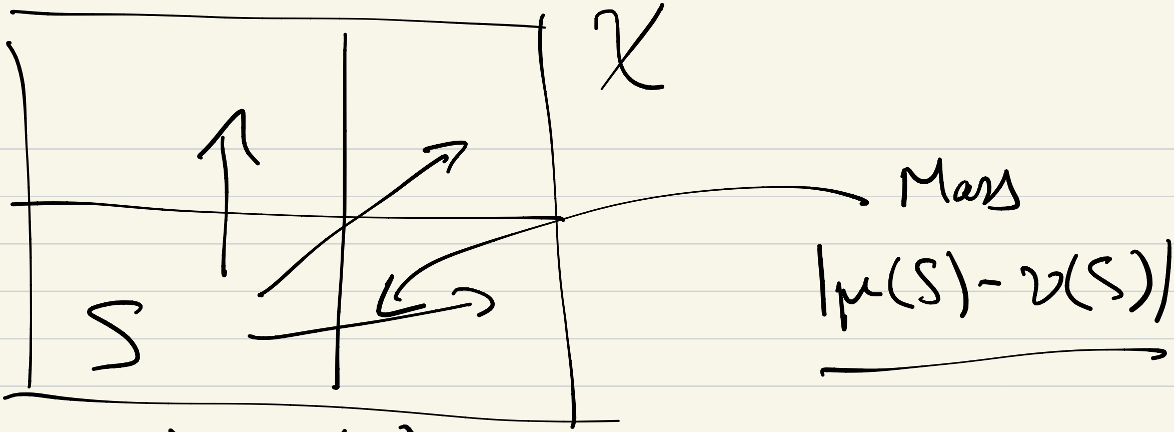
(iii) $\begin{cases} S \in Q_{k+1} \\ S' \in Q_k \end{cases} \Longrightarrow S' \subseteq S$ or $S \cap S' = \emptyset$.

Main Tool: $\mathrm{diam}(\mathcal{X}) \leq 1$

$$\omega_p^p(\mu, \nu) \leq 3^{-k^* p} + \sum_{k=1}^{k^*} 3^{-p(k+1)} \sum_{S \in Q_k} |\mu(S) - \nu(S)|$$

"Dyadic Upper Bound"

μ, ν



$\mu(S) \neq \nu(S)$

Cost of moving mass between blocks:

$$\sum_{S\in Q_k} |\mu(S)-\nu(S)| \underbrace{\cancel{\mathrm{diam}(X)}}_{\le 1}$$

Cost within blocks:

$$\sum_{S\in Q_k} \mathrm{diam}(S)\,\nu(S) \le \quad \mathrm{diam}(S) \le 3^{-k} \quad \forall S\in Q_k$$

Say $k=1$

$$\Longrightarrow \text{Cost } 3^{-1} + \sum_{S\in Q_1} |\mu(S)-\nu(S)|$$

- ε-dimension : $d_\varepsilon(\mathcal{U}) = \dfrac{\log N_\varepsilon(\mathcal{U})}{-\log \varepsilon}$
  $\mathcal{U} \subseteq \mathcal{X}$

- (ε, τ)-covering numbers of μ

$$N_\varepsilon(\mu, \tau) = \inf\{ N_\varepsilon(\mathcal{U}) : \mu(\mathcal{U}) \geq 1-\tau \}$$

- (ε, τ)-dimension of μ :

$$d_\varepsilon(\mu, \tau) = \frac{\log N_\varepsilon(\mu, \tau)}{-\log \varepsilon}$$

- Upper Wasserstein Dimension :

$$d_p(\mu) = \inf\left\{ s > 2p : \limsup_{\varepsilon \to 0} d_\varepsilon(\mu, \varepsilon^\alpha) \leq s \right\}$$

$$\alpha = \frac{sp}{s-2p}$$

Theorem : Let $p \in [1, \infty)$. Then, for all $s > d_p(\mu)$,

$$\mathbb{E}[W_p^p(\mu_n, \mu)] \lesssim n^{-p/s} + n^{-1/2}$$

$$\mathbb{E}[W_1(\mu_n, \mu)] \lesssim n^{-1/d}$$

Proof Sketch.

$\exists \varepsilon' > 0$ such that

$$(*) \qquad d_\varepsilon(\mu, \varepsilon^\alpha) \leq s \quad \forall \varepsilon \leq \varepsilon'$$

Let $k^* = \left\lfloor \frac{\log n}{s \log 3} \right\rfloor$, $k' = \left\lfloor \frac{p}{\alpha} \frac{\log n}{s \log 3} \right\rfloor$

Assume $k^* \geq \left\lceil - \frac{\log \varepsilon'}{\log 3} \right\rceil$

$$\Longrightarrow \varepsilon' \geq 3^{-k^*}.$$

Therefore, in $(*)$, take $\varepsilon = 3^{-k'}$ and get

$$N_{3^{-k'}}\left(\mu, \underline{3^{-\alpha k'}}\right) \leq \underline{3^{k's}}$$

$\exists T \subseteq \mathcal{X}$ s.t. $\underline{\mu(T) \geq 1 - 3^{-\alpha k'}}$ and

$$\underline{N_{3^{-k'}}(T) \leq 3^{k's}}$$

$\exists \{Q_k\}$ s.t.

$$\left|\{S \in Q_k : S \cap T \neq \phi\}\right| \leq \underline{N_{3^{-(k'+1)}}(T)}$$

$$\mathbb{E}\left[W_p^p(\mu_n,\mu)\right]$$

$$\leq 3^{-k^* p} + \sum_{k=1}^{k^*} 3^{-p(k+1)}\, \mathbb{E}\left\{\sum_{S\in\mathcal{Q}_k} |\mu_n(S)-\mu(S)|\right\}$$

$$\Rightarrow \mathbb{E}|\underline{\mu_n(S)}-\mu(S)| \lesssim \sqrt{\frac{\mu(S)}{n}} \wedge \underline{\mu(S)}$$

$$\Longrightarrow \mathbb{E}\left\{\sum_{S\in\mathcal{Q}_k} |\mu_n(S)-\mu(S)|\right\}$$

$$\lesssim 2\underline{(1-\mu(T))} + \sqrt{\frac{|\{S\in\mathcal{Q}_k : S\cap T\neq\emptyset\}|}{n}}$$

$$\leq \underbrace{3^{-k^* p}}_{n^{-p/s}} + \sum_{k=1}^{k'} 3^{-p(k+1)}\left\{\underbrace{2\cdot 3^{-\alpha k'}}_{\text{low order}} + \underbrace{\sqrt{\frac{3^{k's}}{n}}}_{n^{-p/s}}\right\} + n^{-\frac{1}{2}}$$

$$\lesssim \underline{n^{-p/s} + n^{-1/2}}\,,$$